

Gene gating at nuclear pores prevents the formation of R loops

Hélène Gaillard*, Francisco García-Benítez, and Andrés Aguilera*

Centro Andaluz de Biología Molecular y Medicina Regenerativa-CABIMER,
Universidad de Sevilla-CSIC-Universidad Pablo de Olavide, 41092 Seville, Spain

** corresponding authors:*

Andrés Aguilera:

Centro Andaluz de Biología Molecular y Medicina Regenerativa-CABIMER,
CSIC-Universidad Pablo de Olavide-Universidad de Sevilla, 41092 SEVILLE,
Spain, ++34 954 468 372, aguilo@us.es

Hélène Gaillard:

Centro Andaluz de Biología Molecular y Medicina Regenerativa-CABIMER,
CSIC-Universidad Pablo de Olavide-Universidad de Sevilla, 41092 SEVILLE,
Spain, ++34 954 467 728, gaillard@us.es

Keywords: Genome instability, R loop, Transcription, Nuclear pores, Genome organization

Abstract

Transcription is an important source of genetic variability. A large amount of transcription-associated genome variation arises from the unscheduled formation of R loops. We have recently found that physical proximity of chromatin to nuclear pores prevents the formation of pathological R loops during transcription. Our study opens new perspectives to understand genome stability as a function of nuclear location.

Main text

Sequence variations in the form of sporadic mutations or genome rearrangements are necessary for evolution. However, the amount of such genome alterations needs to be kept at low level to maintain genome stability and avoid pathological cell behaviours. High level of genome instability is a hallmark of cancer and other genetic diseases and a comprehensive understanding of the mechanisms leading to genome instability is required for the development of therapeutic and preventive approaches. Beside its crucial role in gene expression, transcription constitutes a major source of genome instability. During transcription, the nascent RNA may hybridize back with the complementary template strand, resulting in an R loop structure formed by the DNA:RNA hybrid and the displaced non-template DNA strand, which remains single-stranded (ssDNA). Although R loops are natural intermediates in a number of physiologically relevant processes, including transcription regulation, mitochondrial DNA replication or class switch recombination of immunoglobulin genes, pathological R-loop accumulation are responsible for a large amount of transcription-associated genome instability (TAGIN). The mechanisms identified so far involve increase susceptibility of the displaced ssDNA to damage, chromatin condensation, interference

with replication and DNA repair pathways. A number of cellular conditions favour the formation of such genome-threatening R loops, such as accumulation of negative supercoiled DNA or defective messenger ribonucleoprotein particle (mRNP) biogenesis, as shown in topoisomerase I and THO mutants in yeast ¹. R loop-mediated TAGIN is emerging as an important contributor to human disease and many gene products involved in R-loop resolution, such as **breast cancer 1 and 2 (BRCA1/2)** and other Fanconi Anemia factors, are linked to cancer predisposition.

Using expression of the human activation-induced cytidine deaminase (AID) as a tool to induce mutations in the displaced ssDNA of R loops, we identified the yeast nuclear basket **Myosin-like protein 1 (Mlp1)** as a factor involved in the prevention of R-loop accumulation ². The absence of Mlp1, its paralog Mlp2 or both proteins leads to AID-dependent hyper-recombination and increased genomic instability. These phenotypes are suppressed by over-expression of RNase H1, which specifically removes the RNA moiety of DNA-RNA hybrids. Direct quantification of DNA-RNA hybrids at different genomic loci using a specific antibody demonstrated that R loops accumulate in *mlp1/2* mutants. Together, these results support the idea that the genomic instability in *mlp1/2* mutants is caused by unscheduled R-loop accumulation. Importantly, such R loops interfere with replication fork progression, thus becoming a major cause of instability. Thus, *mlp1* cells show mild but significant replication defects that can be reverted by RNase H1 overexpression.

The gene gating hypothesis ³ postulates that transient localization of transcribed DNA in the proximity of the nuclear pore complex (NPC) facilitates the formation of an export-competent mRNP. Since Mlp1 is required for gene gating ^{4,5} a pertinent question was whether R-loop accumulation occurred as a consequence of defective gene gating in *mlp1* cells. Indeed, restoration of physical proximity to the NPC is sufficient to

suppress R-loop accumulation in cells lacking Mlp1², as shown with a previously characterised system allowing artificial NPC-tethering of the *GAL* locus⁶. Importantly, artificial gene gating also suppressed R-loop accumulation in a THO mutant, indicating that physical proximity to the NPC generally prevents the accumulation of pathological R loops.

The study indicates that R loops are preferentially formed in the nuclear interior and that gene gating prevents R loop-associated genome instability (Figure 1).

Nonetheless, gene gating may also restrict the freedom of chromosome rotation thus potentially impairing replication fork progression. Recent work has shown that activation of the intra-S checkpoint leads to Mlp1 phosphorylation and gene release from the NPC⁷. Because gene gating is likely limited to highly expressed genes while the entire genome needs to be replicated, replication fork stalling may occur at the nuclear periphery in some cases (e.g. due to torsional stress) and in the nuclear interior in other cases (e.g. due to pathological R loops). Checkpoint-mediated release of gated genes in the nucleoplasm may favour R-loop formation, and these R loops may in turn serve to amplify the checkpoint activation cascade and signal the sites of replication stalling. This hypothesis is consistent with recent work in human cells suggesting that R-loop formation participates in DNA damage signaling and activates the DNA damage kinase *Ataxia telangiectasia mutated* (ATM)⁸.

Gene gating is widely accepted in budding yeast, but whether it is conserved in higher eukaryotes remains yet to be resolved. Both the nuclear basket *nucleoporin 153* (NUP153) and the linker *nucleoporin* NUP93 were shown to interact with transcriptionally active genomic regions at the NPCs in human cells⁹, suggesting that the implication of NPC components in transcription regulation may be a conserved feature in eukaryotes. The Mlp1/2 human homolog *Translocated promoter region* (TPR)

was shown necessary for the formation of nuclear pore-associated heterochromatin exclusion zones in HeLa cells ¹⁰. This raises the possibility that some Mlp1 functions in genome organization and R-loop prevention may be conserved in human cells. Although further work will be required to decipher the role of chromatin spatial dynamics in genome stability, we can conclude that the location of a particular locus within the nucleus is a major determinant of R-loop formation in yeast. These findings add new elements to take into account for a complete understanding of the mechanisms by which R loops are formed and lead to genome instability. Thus, whether the specific sub-nuclear localization is not only a determinant of R-loop formation capability but also to cause genome instability is an intriguing possibility to be tested in the future.

References

1. Santos-Pereira JM, Aguilera A. R loops: new modulators of genome dynamics and function. *Nat Rev Genet* 2015;16(10):583-597. 10.1038/nrg3961.
2. Garcia-Benitez F, Gaillard H, Aguilera A. Physical proximity of chromatin to nuclear pores prevents harmful R loop accumulation contributing to maintain genome stability. *Proc Natl Acad Sci U S A* 2017;114(41):10942-10947. 10.1073/pnas.1707845114.
3. Blobel G. Gene gating: a hypothesis. *Proc Natl Acad Sci U S A* 1985;82(24):8527-8529.
4. Dieppois G, Iglesias N, Stutz F. Cotranscriptional recruitment to the mRNA export receptor Mex67p contributes to nuclear pore anchoring of activated genes. *Mol Cell Biol* 2006;26(21):7858-7870.
5. Tan-Wong SM, Wijayatilake HD, Proudfoot NJ. Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex. *Genes Dev* 2009;23(22):2610-2624. 10.1101/gad.1823209.
6. Texari L, *et al.* The nuclear pore regulates GAL1 gene transcription by controlling the localization of the SUMO protease Ulp1. *Mol Cell* 2013;51(6):807-818. 10.1016/j.molcel.2013.08.047.

7. Bermejo R, *et al.* The replication checkpoint protects fork stability by releasing transcribed genes from nuclear pores. *Cell* 2011;146(2):233-246. 10.1016/j.cell.2011.06.033.
8. Tresini M, *et al.* The core spliceosome as target and effector of non-canonical ATM signalling. *Nature* 2015;523(7558):53-58. 10.1038/nature14512.
9. Ibarra A, Benner C, Tyagi S, Cool J, Hetzer MW. Nucleoporin-mediated regulation of cell identity genes. *Genes Dev* 2016;30(20):2253-2258. 10.1101/gad.287417.116.
10. Krull S, *et al.* Protein Tpr is required for establishing nuclear pore-associated zones of heterochromatin exclusion. *EMBO J* 2010;29(10):1659-1673. 10.1038/emboj.2010.54.

Funding

Research in A.A.'s lab is funded by the European Research Council, the Spanish Ministry of Economy and Competitiveness, and the European Union (FEDER). F.G.-B. is the recipient of a predoctoral Formación de Profesorado Universitario (FPU) grant from the Spanish Ministry of Education, Culture and Sport.

Figure Legend

Figure 1. Possible mechanism of R-loop formation in the nucleoplasm. Physical proximity of transcribed loci to the NPC suppresses R-loop formation in budding yeast. In mutants of the nuclear basket Myosin-like protein 1 (Mlp1) or THO complex, active genes remain in the nucleoplasm and pathological R loops are enhanced. RNAPII, RNA polymerase II.